

## **Cognitive Database: A Step towards Endowing Relational Databases with Artificial Intelligence Capabilities**

We consider *Cognitive Databases*, an approach for transparently enabling Artificial Intelligence (AI) capabilities within relational databases. We first view the structured data source as meaningful *unstructured* text, and then use the text to build an unsupervised neural network model using a Natural Language Processing (NLP) technique called *word embedding*. This model captures the hidden inter-/intra-column relationships between database tokens of different types. For each database token, the model includes a *vector* that encodes contextual semantic relationships. We seamlessly integrate the word embedding model into existing SQL query infrastructure and use it to enable a new class of SQL-based analytics queries called cognitive intelligence (CI) queries. CI queries use the vectors to enable complex queries such as semantic matching, inductive reasoning queries such as analogies, predictive queries using entities not present in a database, and, more generally, using knowledge from external sources. We demonstrate unique capabilities of Cognitive Databases using an Apache Spark based prototype to execute inductive reasoning CI queries over a multi-modal database containing text and images. We believe our first-of-a-kind system exemplifies using AI functionality to endow relational databases with capabilities that were previously very hard to realize in practice.